

# **When Paradigms Meet: Interacting Perspectives on Evaluation in the Non-Profit Sector**

**Abstract:** This study examines a public debate in Australia, arising from a national government report, around how social contribution in the nonprofit sector should be assessed. Guided by several meta-perspectives on evaluation, we identify connections between foundational assumptions and normative positions on evaluation espoused by non-profit organizations (NPOs), and examine the ways in which the inter-paradigmatic context of the non-profit sector contributes to the emergence of NPOs' different normative positions on evaluation. We conclude that particular paradigmatic orientations of NPOs (positivism, interpretivism, constructivism) lead to particular perspectives on how NPOs should engage with alternative paradigms (monism, impartial pluralism, radical pluralism).

**Keywords:** evaluation theory, evaluation paradigm, nonprofit sector, NPO assessment, inter-paradigmatic context

## INTRODUCTION

Controversy about performance measurement is often mistakenly reduced to a mere dispute about technique or information, a position that ignores the significance of the social and political background of measurement processes. Agreement and disagreement about measurement can only be meaningfully understood if consideration is given to the normative positions held by various actors (Hall, 2014) and the paradigmatic assumptions underpinning these views (Guba and Lincoln, 2005). Studies of the influence of new public management within the nonprofit sector (Dart, 2004; Hwang and Powell, 2009; Lewis, 2005) provide empirical evidence of actors who hold different paradigms of evaluation being forced to interact. The case of government outsourced social service delivery to nonprofit organizations (NPOs) (Hwang and Powell, 2009) is a prominent example of social agents with different evaluation paradigms interacting to discuss publicly acknowledged common goals.

Extant theories of evaluation are useful in documenting the particular evaluation approaches and/or instruments that should be used and the normative position that should be adopted about how evaluation should be implemented, as a result of particular foundational premises. For example, evaluators who interpret social reality as objective, social knowledge as discovered, and social intervention outcomes as having a set value will tend to take the normative view that the purpose of evaluation is to discover a single objective truth, usually through measurement as a technical process. Alternatively, evaluators who regard social reality as subjective, social knowledge as constructed, and social intervention outcomes as acquiring different value for different stakeholders will define the purpose of evaluation in terms of a moderating process among different but comparable claims to knowledge made by different social actors (Guba and Lincoln, 1989). However, the connection between foundational assumptions and normative

positions on evaluation requires further investigation, in order to more clearly account for how non-profit organizations (NPOs) deal with competing perspectives on evaluation produced by organizations or institutions from other sectors.

Our subsequent review of theoretical literature on evaluation paradigms highlights the dearth of explanation available regarding the foundational-normative link in evaluation. Additionally, a brief appraisal of empirical studies on social assessment in NPOs identifies the need to examine and develop new understandings of paradigm interactions among social actors. These reviews document the focus of this study on *what foundational assumptions and normative positions on evaluation are espoused by NPO* and *how the inter-paradigmatic context of the NPO sector contributes to the emergence of different normative positions on evaluation*.

In order to address these questions we study a public debate between an Australian Federal Government agency and Australian NPOs on the question of how the social contribution of NPOs *should* be evaluated. Using this rare opportunity of a real-life example of inter-paradigm relations occurring among different social agents where a key normative question about evaluation is publicly asked, the study contributes new insights into the different paradigmatic preferences of government and NPOs. We propose that knowledge of the different foundational assumptions and normative positions espoused by NPOs in an inter-paradigmatic context could assist in the formulation of more comprehensive and nuanced theories and practices of evaluation and inter-sector communication.

## **EVALUATION PARADIGMS: THEORETICAL AND EMPIRICAL RELEVANCE**

Foundational assumptions that underpin various normative positions have led to different forms of evaluation (Owen, 2006), different rationales of evaluation (Alkin, 2012) and different evaluation logics (Hall, 2014). Indeed, the “simple” act of defining evaluation points to the significance of paradigmatic distinctions in this context. Compare, for example, Alkin’s widely cited definition of evaluation as an “application of the repertory of social research methods to provide credible information that can aid in the formation of public policy, in the design of programs, and in the assessment of the effectiveness and efficiency of social policies and social programs” (Alkin, 2004, p. 217) with Guba and Lincoln’s constructivist understanding of evaluation as a social process of negotiating values among different social actors through a ‘hermeneutic dialectic’ of creating – rather than discovering – plural and subjective ‘realities’ (Guba and Lincoln, 1989, p. 17). Meta-perspectives on evaluation emphasize and explore such paradigmatic distinctions.

We therefore explore the evaluation theory literature in order to document the logical connections to be made between foundational assumptions and normative positions in evaluation. We then examine empirical studies on NPO assessment to highlight contexts of paradigm interaction and the practical relevance of studying the communications and arguments of social actors (such as NPOs and government) who are drawn to interact while upholding different paradigms of evaluation.

### ***Theories of Evaluation and Paradigmatic Distinctions***

In order to explore the link between foundational assumptions and normative positions, we discuss three meta-perspectives that employ paradigm-defining foundational distinctions. We compare and integrate Guba and Lincoln (1989) and Mertens (1999, 2009) as representative of

this discussion in the broader evaluation literature, and include Hall (2014) to document paradigmatic distinctions that tend to operate in evaluation logics specific to the nonprofit sector (see Table 1).

Insert Table 1 here

Using a historical lens, Guba and Lincoln (1989) identify four types of evaluation, namely, measurement, description, judgment, and constructivism. Within the measurement generation, the evaluator has a technical role, to provide measurement instruments for any aspects deemed worthy of evaluation. On the other hand, description focuses on identifying strengths and weaknesses relative to pre-set objectives, while judgment emphasizes the need to evaluate objectives as well, but relative to pre-set standards. These first three generations of evaluation are critiqued as facets of a scientific paradigm, and it is in contrast with this paradigm that the authors develop their preferred alternative, namely responsive constructivism. The latter is opposed to the scientific paradigm in three crucial respects. First, it rejects the managerialism that dominates the previous three generations, as it challenges the hegemony of donor-funder interests and objectives in the social agenda setting processes. Second, it accuses these views of insensitivity to value pluralism and, third, it challenges their propensity toward positivist forms of inquiry. Instead, responsive constructivism regards evaluation as a primarily political process, more specifically as a response to issues and claims, involving stakeholders with different interests and powers, and promoting an agenda for negotiating these differences through a fairer and more egalitarian process of inquiry (Guba & Lincoln, 1989). The responsive constructivist argument can thus be summarized as follows: because social experiences are different and irreducible to one another, these differences should be adequately voiced and reflected in evaluations. This argument illustrates the logical connection between a foundational premise

about the nature of social experience and a normative conclusion about the purpose of evaluation and how it should be conducted.

Mertens (1999, 2009) opposes interpretive evaluations to post-positivist evaluations, by criticizing the latter's assumptions of objective reality, of single, knowable truth, and their confidence in probability calculations. However, while acknowledging interpretive evaluations as progressive in their recognition of 'realities' as socially constructed, therefore plural and irreducible to one another, she contends that the presupposed impartiality of interpretivism, anchored in the notion that no social perspective is 'truer' than another, has the effect of enforcing the status quo and excluding, or further marginalizing, the vulnerable. Concerned with inclusion and empowerment, hence with the role of the evaluator as challenging the existing order and advocating for disadvantaged groups, Mertens promotes a transformative paradigm which encourages inclusive evaluation (Mertens, 1999; Mertens, 2009). The transformative paradigm, which aligns with Guba and Lincoln's responsive constructivism (as shown in Table 1), is particularly distinctive due to its foundational premise about the centrality of meaningful social change (towards justice and human rights) to evaluation processes. Consequently, Mertens advances evaluation as a tool to confront social inequality and facilitate equality (Mertens, 1999; Mertens, 2009; Alkin, 2012). Her argument provides an example of foundational premise about the nature of the evaluation process leading to a normative conclusion about the processes and effects involved in conducting evaluation.

Finally, the concept of 'evaluation logic' has often been used to identify stable and predictable connections between foundational premises about social reality and normative approaches to

evaluation (Scriven, 1980; Stufflebeam, 2001; Owen, 2006). Hall's (2014) account of evaluation logics illustrates such connections in the context of the nonprofit sector. Following analysis of evaluation methods commonly employed in this sector, he identifies three types of evaluation logic – namely, scientific, bureaucratic, and learning logics. Accordingly, scientific logics assume that emphasis is on proof; data collection processes and analyses are objective; disagreements on the evaluation process are contingent and can be avoided by improving the methods and instruments; the complexity of social phenomena should be accounted for using simpler, essentializing frameworks; and the evaluator is a scientist who 'conducts research and reports study findings' (Hall, 2014, p. 325). On the other hand, bureaucratic logics highlight categorization (rather than evidence gathering) as the central evaluation activity, assume the evaluation process to be linear or sequential, focus on intended effects, and regard the evaluator as an implementer of plans and objectives. Finally, learning logics emphasize data richness, rely on more egalitarian assumptions about professional skills and expertise, which favour more transparent and democratic views of knowledge, regard the main purpose of evaluation to be one of belief revision, and assume the evaluator to be a 'facilitator' of other stakeholders' involvement in the evaluation process (Hall, 2014). Here we can infer paradigmatic distinctions that can be made among these logics, based on identified connections between foundational premises and normative positions. For example, in the case of scientific logics, confidence in the possibility of objective knowledge leads to the prescription that evaluation should be evidence and measurement based. In bureaucratic logics, focus on values such as utility and efficiency legitimizes centralized evaluation approaches based on categorizing and structuring data through abstraction and functional manipulation. In contrast, learning logics stem from foundational assumptions of pluralism and unpredictability of social phenomena, hence belief revision

prominently guides evaluation approaches in this perspective. Approaches centred on learning logics tend to be more communicative, participative, iterative and reliant on feedback loops.

By examining the underlying assumptions across these meta-theoretical conceptions of evaluation, three important ideas about foundational-normative connections arise. First, scientific foundational premises will naturally lead to normative positions that support measurement, unifying frameworks and (at least) some shared and pre-set standards, while constructivist premises will favour normative positions that emphasize value pluralism and political negotiation of desirable outcomes by different social actors (Guba & Lincoln, 1989). Second, approaching the evaluation process as ‘neutral’ (whether informed by scientific objectivity or by recognition of plural values) leads to passive normative positions supporting the status quo and perpetuating its exclusive and marginalizing effects, while transformative evaluation, focusing on the need to induce social change, attracts normative positions which centre their support on the interests of the powerless (Mertens, 1999, 2009). Third, evaluation logics based on scientific and bureaucratic foundational premises, focusing on objective evidence and systemic efficiency, privilege the evaluator’s role as one of expert appraisal and decision making, while learning premises, open to feedback and revision in light of changes in social goals and values, assume the evaluator as a facilitator of equitable stakeholder participation in evaluation deliberations (Hall, 2014).

### ***Inter-Paradigmatic Contexts in NPO Assessment***

While the theoretical connections between foundational premises and normative positions argued by these three meta-perspectives are compelling, we are yet to consider what happens with these



arguments in practice, when different paradigms meet. We define an inter-paradigmatic context as a situation in which social actors adopting different paradigms of evaluation must engage, for practical reasons, in projects that involve the evaluation of social interventions and also necessitate some form of agreement on how this evaluation should be carried out. We are particularly interested in the challenges encountered by social agents like NPOs when having to interact with powerful others, such as government, who may approach evaluation from a different paradigm.

The question of how the social services and programs of NPOs should be evaluated has become an increasingly public question, due to accountability pressures from several sources: donors and sponsors; government; the public in general (Carman, 2007; Cheverton, 2007; Ebrahim and Weisband, 2007; Eckerd and Moulton, 2011; Quarter and Richmond, 2001; Thomson, 2011); and, importantly, from NPOs themselves, as they wish to identify the best approaches to managing their own activities (Gugerty, 2009). However, despite wide cross-sector agreement on the importance of this question, deep-seated disagreements and ambiguities around how these evaluations should be conducted persist (Brown and Gaughlin, 2009; Herman and Renz, 1999; Polonsky and Landreth Grau, 2008).

Empirical studies of evaluation processes in NPOs have found that increasing external accountability pressures experienced by these organizations, under the influence of new public management trends, have led to the adoption by NPOs of centralized and standardized measurement frameworks and practices for reasons of compliance rather than organizational learning or strategic development of social programs (Thomson, 2011). As a consequence, these frameworks have diluted rather than enhanced the service improvement abilities of NPOs

(Carman, 2007; Carman and Fredericks, 2008). These compliance and accountability pressures are often articulated in terms of the imperative to be ‘business-like’ (Dart, 2004, p. 290), which affects four important areas of NPO activity: goal setting, service delivery processes, organizational management, and organizational rhetoric (Dart, 2004). A recurrent aspect of being ‘business-like’ is emphasis on measurement as the central approach to evaluation (Murray, 2005; Herman and Renz, 2008; Sowa et al., 2004).

However, empirical research into the diversity of purposes, which characterizes the non-profit sector, questions the usefulness of both centralization and measurement when it comes to evaluation in, and of, NPOs (Eckerd and Moulton, 2011). Based on this evidence, Moulton and Eckerd (2012) argue in favour of public funding of NPOs, to avoid goal capture by business-oriented perspectives – and, at the same time, against centralization and measurement, to preserve NPOs’ diversity of approaches to their public roles. Furthermore, evidence of the effects of the managerialization and rationalization of evaluation processes and outcomes in NPOs (Frumkin, 2002; Hwang and Powell, 2009) indicates that an increase in compliant standardization is achieved at the expense of experimentation, which lies at the centre of NPOs’ unique capacity for social innovation.

After summarizing the extant empirical literature, Herman and Renz (1999, 2008) develop and formulate several theses regarding NPO organizational effectiveness (understood as NPOs’ ability to reach their set goals), which include, *interalia*: NPO effectiveness is multi-dimensional and cannot be reduced to a single measure; NPO effectiveness is a social construct; and, program outcome indicators are limited in their ability to reflect NPO effectiveness, due to the changing nature of social goals (Herman and Renz, 2008).

While the conclusions drawn by the above studies are useful, they are also reflective of the researchers' foundational assumptions and normative positions on evaluation in NPOs. Little is revealed, in a systematic way, about what NPOs themselves actually think about the nature of evaluation in the non-profit sector and how it should be carried out. This lack of voice is understandable, given the increasing compliance and accountability environment these organizations operate in.

We note that analyses of the multiple influences of bureaucratic rationalization and business managerialism on NPO activity (e.g. Hwang and Powell, 2009) emphasize the inter-paradigmatic contexts and tensions created, as a result, by these influences. Using Hall's (2014) typology of evaluation logic, for instance, we interpret the empirical research on NPO assessment cited here to suggest that, while donors, funders and sponsors (whether from government or the private sector) may have their perspectives on evaluation shaped by scientific and/or bureaucratic logics, these expectations may come into conflict with the learning logics traditionally adopted by NPOs. When these paradigms clash, the conclusion of the above empirical studies seems to be that NPOs will inevitably comply and yield to the outside pressure, despite awareness of the limiting and distorting effects of this pressure on NPO activities and outcomes. We argue that more research into what NPOs themselves have to say about the nature of evaluation and its processes is likely to reveal a more nuanced picture of inter-paradigmatic relations. We therefore pose the questions: *What foundational assumptions and normative positions on evaluation are espoused by NPOs? And, how does the inter-paradigmatic context of the NPO sector contribute to the emergence of NPO's different normative positions on evaluation?*

To address these questions, we investigate normative views on NPO evaluation, as expressed by key actors such as NPO practitioners and government policy advisers, who represent a crucial source of information for understanding paradigmatic distinctions. A public inquiry undertaken by the Australian government in 2009-2010 into the contribution of the NPO sector, in which NPO's respond to the centralized measurement framework proposed by government, constituted a unique opportunity for the NPOs to directly and explicitly engage in inter-sector communications about NPO evaluation. Hence, we analyze the Productivity Commission's 2010 Final Report and 54 NPO submissions influencing its compilation and content. The focus of our investigation is on these actors' normative views, namely, how they believe NPO evaluation should be undertaken, as well as their explicit or implicit foundational assumptions.

In examining the texts of these submissions, we identify the foundational premises used by the NPOs to explain and justify their normative positions relative to the proposed introduction of a common measurement framework (CMF). We then classify and further analyze the types of foundational-normative arguments made, in order to understand what considerations may influence the NPO's normative responses in this inter-paradigmatic context.

## **CONTEXT OF OUR STUDY**

The most recent *Contribution of the Not-for-Profit Sector* (2010) enquiry by the Productivity Commission builds on earlier attempts at NPO sector evaluations by the Australian Government in 1995, through its Industry Commission report *Charitable Organizations in Australia* (Industry Commission, 1995) and its later *Report of the Inquiry into the Definition of Charities and Related Organisations* (Commonwealth of Australia, 2001), aimed at facilitating a common understanding of the social contribution of the NPO sector. In commissioning the enquiry, the

then Australian Labor Party (Labor) government openly acknowledged tensions in government sector relations with NPOs and professed willingness to progress a cooperation agenda relevant to program management and evaluation. It was also the first time that public submissions were requested on the very question of NPO social contribution evaluation. The initiative itself was widely welcomed as an attempt by government to spur public debate around issues of fundamental concern for social policy and social services management (Housego and O'Brien, 2012).

While the Productivity Commission undertook a broader consultation exercise, which included tax treatment and volunteer involvement among its topics, part of this exercise was the Commission's proposal for a Common Measurement Framework (CMF), which was regarded as a research-based instrument for increasing the accountability of the NPO sector, for centralizing its performance data, and for shifting measurement emphasis from inputs and outputs to outcomes and impacts (Productivity Commission, 2010, pp. xxvi, xxx, 34). The CMF proposal attracted significant interest and diverse reactions from the NPO sector. In an initial foray of the submissions, we noted that the responding organizations reflected the composition of the Australian NPO sector (Industry Commission, 1995; Productivity Commission, 2010); that is, they were primarily engaged in community or health service provision, member services (including peak bodies, advocacy, research, training and support services), philanthropy, or combinations thereof<sup>1</sup>.

---

<sup>1</sup> This categorization is typical of literature regarding the NPO sector, and was also used in the enquiry process.

The Commission's Report (referred to hereafter as the Report) prefaced its CMF proposal with an explicit recognition of the nonprofit sector's specific focus on community purpose (p. xxv) and potential for social innovation (p. 16). In doing so, it echoed the acknowledgments made by previous government reports (e.g., Industry Commission, 1995). However, the 2010 Report critiqued past and current practices by Federal and State government agencies which regard NPOs as 'a cheap way of providing social services', which is likely to 'undermine the sustainability' of NPO activities (p. 384).

Due to the unprecedented openness of government to this debate, we expected that this important episode in the history of Australian Government-NPO relations, which publicly involved all sectors of society, would trigger strong public positions on the part of NPOs, with rich arguments that would clearly reveal sources of agreement as well as disagreement between them and government. Hence, we considered it an ideal context for investigating a more extensive range of foundational assumptions and normative positions than would otherwise be available. For many NPOs, this public debate was a rare opportunity to reflect more broadly and carefully on the foundational premises of their work, and to articulate foundational-normative arguments in ways that *ad hoc* field enquiries from academic researchers may not be able to elicit.

## **RESEARCH METHOD AND PROCESS**

Using the 2009-2010 public debate as a case study of a potentially inter-paradigmatic context, we sought to address the research questions regarding the foundational assumptions and normative

positions espoused by the NPOs, as well as the contribution of the inter-paradigmatic context of the NPO sector to the emergence of different positions on evaluation.

Our first step was to collate the central document (Productivity Commission, 2010) and the associated 319 submissions generated by the preceding enquiry<sup>2</sup>. We undertook purposeful sampling based on conceptually derived criteria (Creswell, 2013) to identify submissions by active NPOs which took an explicit normative position on how NPOs' social contribution should be evaluated. Hence, the following inclusion criteria were used: 1) the organization authoring the submission was confirmed to be an Australian nonprofit entity<sup>3</sup>, active at the time of study, and engaged in public reporting of their own services and/or programs; and 2) answers were provided to the Report's questions about how NPO social contribution should be evaluated and whether the CMF (as an example of evaluation based on centralized measurement) should be adopted or not<sup>4</sup>. The final selection included N=54 submissions. This purposeful sample was not representative of the NPO sector as a whole but of NPOs that chose to take a normative position on evaluation of the NPO sector (i.e. chose to address the evaluation question by making specific recommendations on how it should be carried out).

We then read through these 54 documents (approximately 900 A4-size pages of text), and selected from this text examples of foundational assumptions on evaluation and normative

---

<sup>2</sup> As of 14 March 2015, all submissions referred to in this study are publicly accessible at <http://www.pc.gov.au/inquiries/completed/not-for-profit/submissions>.

<sup>3</sup> Submissions from entities such as government agencies, trade unions, religious institutions, private sector consultancies, academic researchers and private individuals were excluded.

<sup>4</sup> Submissions which did not include terms specific to evaluation such as "social contribution" and "common measurement framework" (or "CMF"), that is only responded to other questions raised by the Report e.g. taxation of NPOs, reduction of red tape and volunteer involvement, were excluded.

positions relative to the CMF. The examples were tabled and progressively grouped into themes which emerged as the reading and selection progressed. Two researchers undertook the analysis independently of each other. The researchers discussed their differences of interpretation on a continuous basis, to achieve consistency in the interpretation and attribution of foundational premises and normative positions. These discussions were aimed at reaching consensus on the findings but also increasing our depth of interpretation. The few disagreements identified during this process were resolved by grouping narrower subthemes into broader themes – until both researchers agreed on the three final themes and the examples/NPOs grouped under each theme. In sum, to enhance the trustworthiness of our analysis – in particular the credibility (clear identification and description of participants), dependability (stability of the data), conformability (congruence of two or more independent researchers) (Lincoln and Guba, 1985) , we used multiple rounds of discussions that refined the researchers’ common understanding of the data.

### **FINDINGS: FOUNDATIONAL ARGUMENTS**

The CMF proposed and advocated in the Report anchors the question of how the social contribution of NPOs should be evaluated, and thus provides expression of the Commission’s normative position. Correspondingly, we have found that the NPOs’ normative positions, as expressed in the submissions, were formulated in relation to the CMF.

In our analysis we found that foundational assumptions were most explicitly presented by NPOs to express disagreement with specific points in the CMF proposal. In comparison, foundational assumptions were not always clearly stated where the NPO’s position was more nuanced, while in cases of extensive agreement with the CMF proposal such assumptions were implicit. Hence, it is in the manifestation of disagreement that we found the richest data on foundational



assumptions. According to this pattern, where the NPO shared the same paradigm of evaluation as the CMF proposal the discussion was confined to technical aspects of how centralized measurement should be planned and implemented. In such cases, explicit reflection on the foundational assumptions of the shared paradigm would have been considered unnecessary. In contrast, such reflection became essential for those NPOs who approached evaluation from a different paradigm, as they needed to justify their disagreement with the CMF proposal.

In sum, we have found three distinct arguments employed by disagreeing NPOs to explain their opposition to the CMF by reference to their different foundational assumptions: non-measurability, contextualization, and emergence. To present each of these foundational arguments and the broader themes they indicate, we start with the normative position evidenced by the Report followed by examples of a range of normative positions identified in NPO submissions from strong disagreement to strong agreement. As explained above, data on foundational assumptions becomes progressively less explicit as agreement strengthens.

### ***The Non-Measurability Argument***

Notwithstanding acknowledgement of the complexity of causal relationships between social phenomena, data collection difficulties and inherent limitations to measurability for certain types of outcomes and impacts, the Report justifies the need to use measures in terms of improving “the allocation of resources” as well as “efficiency and effectiveness”, overseeing the outcomes of large scale policy changes, and increasing NPO accountability to donors and the public (Productivity Commission, 2010, p. xxvi). Interpreted as an assumption of measurability, this aspect of the CMF is criticized by some NPOs (n = 10) on the grounds that, due to their

qualitative nature, certain values that are central to NPO activity cannot be meaningfully measured<sup>5</sup>:

While philosophers and legal theorists might argue about whether everything of value can be reduced to a... metric (e.g. the satisfaction of individual preferences), the majority of ordinary people do not think about value in these terms. Rather than being concerned solely with efficiency, most people are also concerned with issues such as fairness. (CSSA<sup>6</sup>, #117, p. 7)

Social services make a moral and ethical contribution to society... Moral and ethical principles such as fairness, equality and respect for human dignity cannot be fully captured by an economic approach to valuation. Their value is not exhausted by their contribution to social efficiency. (CSSA, #117, p. 9)

It was observed, for example, that non-measurable aspects of social value, such as social inclusion and community cohesion (Health and Community Services Workforce Council (HCSWC)), were often more important than measurable ones (Australian Council for International Development (ACFID), Cancer Council Australia, Peakcare Queensland). Consequently, against what the Report recommends and the CMF illustrates, these NPOs argued that the number of quantitative measures in an evaluation framework should not be increased (HCSWC), as in their experience this did not necessarily lead to an increase in the ability to capture the full range of outcomes delivered (Victorian Council of Social Service (VCOSS)). It was also suggested that longer-term social objectives, which the Report itself agrees should be prioritized over short-term ones (Productivity Commission, 2010, p. 34), are also significantly more difficult to measure compared to short-term goals; therefore the idea that the most important aspects of evaluation lie in measurement should be abandoned (Illawarra Forum, Kids Under Cover).

---

<sup>5</sup> To simplify the presentation of our findings, indicative examples (rather than complete lists) of submissions are included to support each point.

<sup>6</sup> All NPO names longer than three words have been abbreviated throughout the text. Their full names and corresponding submission numbers are listed in Table 3.

Other, more moderate, responses (n = 11) also employed elements of the non-measurability argument. For example, some NPOs stated that, while quantitative methods may be meaningful at sectoral (macro) level, their relevance is considerably reduced at the level of field experiences of social interventions (the micro level). In response to recommendations that advocated the CMF as an impact measurement tool across the sector (Productivity Commission, 2010, pp. xli-xlii), these NPOs acknowledged the rationale of the CMF as valid from the government's perspective but encouraged the use of qualitative techniques alongside quantitative measures (Australian Council of Social Service (ACOSS), Communities@Work, Vision Australia). Raising concerns that the CMF displayed a predominantly quantitative indicators-based structure, these submissions expressed a strong preference for the use of qualitative techniques.

There were also NPOs (n = 7) that embraced the idea of measurement and shared the Commission's confidence in this approach (e.g., Ability Options, PilchConnect). These NPOs largely supported the adoption of the CMF, welcoming the improvements in transparency, accountability and efficiency that it would bring.

### ***The Contextualization Argument***

The Report validates the need to have a centralized form of evaluation for the whole NPO sector, in terms of improving society's understanding of the role and contribution of NPOs, and simplifying and reducing national accountability and reporting requirements (Productivity Commission, 2010, p. xxiii). Consequently, the CMF is proposed in the Report as an instrument of this centralization.

In reaction to this proposal, a number of submissions (n = 19) argued for the fundamental importance of context to any attempts to evaluate social service contribution. They maintained that, since social phenomena experiences and values are diverse and their diversity is irreducible to a common denominator, evaluated outcomes cannot be meaningfully centralized:

NFP services are in the best position to design, deliver and adapt services according to client and community needs. It is essential that government recognise the importance of local knowledge and the appropriateness of service delivery in a local context. (Western Australian Network of Alcohol and other Drug Agencies (WANADA, p. 3)

[What should be encouraged at national level is] the development of local projects to meet community needs that occur from the ground up (AGPN, p. 8). Not for profits need opportunities to be innovative to meet local needs (AGPN, p. 9).

A[n]... important obstacle [to centralised social evaluation] is *incommensurability*. Goods such as social justice or human dignity cannot be weighed up against each other using a common metric (CSSA, p.8)... The most efficient and effective way to promote the wellbeing of disadvantaged individuals and communities depends on the individuals' circumstances and on local conditions. When focused on impacts and allowed sufficient autonomy community based agencies can harness local knowledge to make the best use of their scarce resources (CSSA, p. 13).

Vocal advocates of this position were Anglicare Australia, AGPN, CSSA and WANADA. These NPOs were concerned that a centralized framework would not be context-sensitive enough to adequately measure their work, as they feared that homogeneous measures would discount the value created through the inherent diversity and innovative capabilities of the nonprofit sector.

Based on the contention that meaningful general indicators are difficult to design for such a diverse and complex sector, questions were raised as to whether the introduction of centralized indicators may reduce contextual responsiveness to local needs and disadvantage smaller community-centered organizations (AGPN). Hence different solutions were suggested by different NPOs: (1) if a CMF-type of instrument is to be introduced, then smaller organizations should be exempted from centralized reporting (The Smith Family); (2) CMF should be

combined with more flexible and open-ended evaluation systems, such as Sen's capabilities approach (Sen, 1999) to produce context-sensitive need assessment outcomes (Anglicare Australia); (3) other frameworks (e.g., the results-based accountability framework), already used by some large NPOs, would be more appropriate to implement instead of the CMF, to avoid the pitfalls of a 'one-size-fits-all' approach (New South Wales Family Services (FamS)<sup>7</sup>, UnitingCare Australia); and (4) due to the nonprofit sector's distinctive focus on social intervention where government and business sectors do not reach, NPOs in general should be allowed to develop their own different evaluation approaches, which are likely to better reflect this focus (Local Community Services Association (LCSA), WANADA).

Moderate positions on this issue (n = 12) were more receptive to the idea of centralizing evaluation systems, as they considered the CMF to be useful in assisting national policy objectives, such as social inclusion, regional development, and national well-being – in broad programs which involve a long-term policy horizon, both for government and for the NPO sector (Australian Red Cross). They did, however, clearly separate evaluation needs at macro (sectoral) level and similar needs at micro (organizational) level. While the CMF was accepted as appropriate and legitimate for the former, doubts were raised about its application for the latter. These NPOs stated that grassroots, practitioner-centered knowledge should be paramount in creating the most adequate forms of evaluation (The Benevolent Society) and that government should refer to existing assessment frameworks before launching into the adoption of new frameworks (ACOSS).

---

<sup>7</sup> This is the NPO's own preferred abbreviation, widely accepted in the Australian community.

A number of NPOs (n = 9) welcomed the CMF initiative. They valued it as an opportunity to integrate and coordinate approaches to disparate local needs and achieve efficiencies in macro-funding systems. However, they also appreciated that it was much harder to collect data at macro-level, on national policy issues, than at micro-level, on local programs (e.g. The Australian Lung Foundation).

### ***The Emergence Argument***

The Report advocates for the need to make predictions and assumptions about the future, on the grounds that, if longer-term outcomes and impacts of social services are correctly anticipated, the likelihood of agreement between government and NPOs on evaluation processes will increase – all to the benefit of ‘the well-being of society’ (p. iv). Hence, the public and stakeholder demand for greater accountability is invoked to justify the need to use pre-set objectives and plan program implementation (Productivity Commission, 2010, pp. xxx, 49). In support of all these recommendations, the Report promotes the CMF as a refined means to emphasize outcomes and impacts over inputs and outputs – a declared point of difference from past government and NPO evaluation practices.

Resisting the assumption of design and planning, another recurrent argument presented by NPOs (n = 24) was that, since social experiences are in a constant state of flux, *a priori* assumptions of objectives and outcomes may be very limiting. In other words, since social phenomena and experiences are emergent and continuously changing, planning outcomes of social interventions based on pre-set objectives may have limited value. The most prominent exponents of this position were Anglicare Australia, FamS, LCSA and UnitingCare Australia. Examples include:

Evidence [of social needs changing in time]... is [also] empirical in that it is grounded in observation and interpretation. But though descriptive, it is primarily concerned with meaning, understanding and, finally, moral value which cannot be reduced to predetermined outputs or benchmarks... As a general rule, the more vulnerable the client, the more qualitatively-based... the service and outcome. (Anglicare, p. 11)

Services like ours are changing constantly to meet local needs... (LCSA, p. 13)

Many organisations in the NFP sector [are] small community-focussed entities which expand and develop organically to meet changing needs and aspirations of members and service users. (UnitingCare Australia, p. 8)

It was argued, for instance, that using measures of pre-defined outcomes and impacts may become inadequate if new social conditions emerge as a result of complex factors (Anglicare Australia) and if longer-term sustainable outcomes only become visible years later than initially expected (Australian Red Cross). To allow for emergent (rather than planned) outcomes, use of the RBA framework was proposed (FamS, LCSA, UnitingCare Australia), and successful examples of accounting for emergent social needs were invoked. It was suggested that *principles* should be used instead of a *framework*, to facilitate the development of meaningful evaluation instruments at grassroots level, premised on the belief that general guidelines and principles at sectoral (macro) level may be better suited to allow objectives to evolve with the circumstances of the social need or problem addressed at organizational (micro) level (Illawarra Forum).

There were also a number of NPOs (n=10) which visibly shared the Report's perspective that the CMF would shift emphasis from inputs and outputs to outcomes and impacts and that, through planning the latter, it would effectively support the distinctive social role of NPOs to respond to social needs and problems in ways that are meaningful to the social services users (Productivity Commission, 2010, p. 32). Representative of this position were Ability Options and the South Australian Council of Social Services (SACOSS).

## ANALYSIS: NPO GROUPS

Taking into account the three arguments presented above, we identified three major themes, reflecting tensions between two opposite attributes: measurability vs non-measurability, centralization vs contextualization, and design vs emergence. We noted that the NPOs' positions tended to cluster towards one or the other term of each dichotomy – in the case of the first two themes, with a small number of moderate positions in between.

This grouping, presented in Tables 2 and 3, enabled us to observe the coupling of the first two themes: those submissions which favoured measurability also favoured centralization, while those questioning measurability also questioned centralization. Table 2 illustrates with relevant quotes the foundational assumptions and normative positions characterizing each group, while Table 3 lists the NPOs categorized under each group.

Insert Table 2 here

Insert Table 3 here

As shown in Tables 2 and 3, we identify three distinct groups of submissions. The first group (G1) is represented by NPOs that shared the Report's assumptions of measurability, centralization, and design. The second group (G2) consists of NPOs that considered that social interventions, inherently dynamic and unpredictable, could not, strictly speaking, be planned. However, they still saw merit in measurement and centralization, at least at broader, sector level. Finally, NPOs in the third group (G3) were critical of the CMF on all three accounts, as they adopted the three foundational arguments presented above.



After categorizing the NPO submissions based on their foundational assumptions, we then examined the normative positions adopted by these submissions relative to how evaluations should be conducted and the role of the CMF in this process. These results are also included in Tables 2 and 3. We found that the G1 submissions largely supported the CMF, as they agreed with its rationale on all three accounts: measurability, centralization, and design. The few cases where amendments were requested referred to improvements in data sharing (SACOSS) and increases in government support (Alcohol and Other Drugs Council of Australia (ADCA)) – operational aspects which did not affect the paradigmatic positions of these NPOs. Common to these submissions was the view that alignment between their own perspective on evaluation and the paradigm supporting the CMF is fundamentally unproblematic: any difficulties would be technical in nature, to be resolved through improvements in measurement.

On the other hand, G2 and G3 submissions identified specific paradigmatic differences between their perspectives and the CMF. This enabled us to summarize an important distinction made by NPOs between their purposes and those of government in evaluation processes. This distinction can be explained in terms of different levels of analysis. Thus, NPOs in these two groups positioned their own evaluation perspective at a micro level, characterized by qualitative experience of social interventions, local context and the local expertise of the evaluator. In contrast, they interpreted the perspective espoused by the CMF to be reflective of a macro level, characterized by the aggregation of social intervention experiences into metrics, the centralization of similarities across the non-profit sector, and the generalizing expertise of the evaluator.

Accordingly, the G2 submissions accepted measurement and centralization as justified at macro level and agreed with the introduction of the CMF but emphasized this rationale as distinct and

separate from that of organizational (micro level) evaluations (Australian Red Cross).

Consequently, they insisted that appropriate government funding and infrastructure should be provided to support this separate effort (The Benevolent Society, The Wilderness Society), which is otherwise likely to hinder the work of NPOs and distract them from their very different purpose (Communities@Work). Invoking the different rationales driving the two levels of analysis, some submissions also insisted that small NPOs should be exempt from macro level reporting (Australian Red Cross, The Wilderness Society).

On the other hand, G3 submissions had difficulties accepting the CMF in its proposed form. The recurrent reasons invoked clearly related to their disagreement with the measurability, centralization and design premises of the CMF. We found, however, three distinct approaches to expressing their disagreement, which led us to dividing this group further, into G3a, G3b and G3c. These three subgroups are also documented in Tables 2 and 3.

Thus, G3a submissions propose foundational changes to the CMF, to allow for experiential narratives rather than measurement (Illawarra Forum), grassroots autonomy of evaluation processes rather than centralization (Anglicare Australia, Illawarra Forum), and emergent outcomes rather than pre-set objectives (Australian Research Alliance for Children and Youth (ARACY), The Smith Family, UnitingCare Australia). One step more challenging, G3b submissions suggest the replacement of the CMF with other frameworks, for two main reasons: first, that given the distinct nature of social work, a common measurement framework for evaluation in the NPO sector should allow for more flexibility and openness in its foundational premises (Cancer Council Australia); and second, that the government should refer to the practitioner-based expertise of NPOs that have already developed and employed more suitable

frameworks for quite a while (FamS). Finally, G3c submissions did not accept the CMF and found it of limited relevance for supporting the work of NPOs (AGPN, CSSA, LCSA, WANADA). The submissions in this group presented the most articulate and developed foundational arguments, which led to the three arguments and corresponding themes discussed above and formed the basis of our analysis.

## **DISCUSSION**

Our analysis illustrates the interplay between foundational arguments and normative positions in shaping NPO perspectives on evaluation. This enables us to categorize the NPOs under study according to their foundational assumptions relative to the three themes and link these with their normative positions relative to the CMF. But further connections can also be made, in two directions. First, the different evaluation perspectives of the three groups can be interpreted to reflect different paradigms of evaluation as defined in the theoretical literature and summarized in Table 1. Secondly, these different paradigmatic orientations of the three groups also explain the NPOs' different perspectives on the inter-paradigmatic context under study – in other words, the way they understand, react and choose to deal with perspectives on evaluation entertained by other social actors (in this case, the perspective reflected in the CMF). Both types of connections, namely with paradigmatic orientations and perspectives on the inter-paradigmatic context, are documented in Table 4.

Insert Table 4 here

Accordingly, we can identify G1 NPOs as reflecting a positivist paradigm. Their response to the CMF proposal demonstrates strong confidence in measurement as a value neutral process, their justification of centralization relies on the assumption of a single locus of expertise, and their confidence in planning and pre-set objectives suggests the existence of a single objective truth accessible to the expert. But this form of response reveals not only the paradigmatic orientation of G1 NPOs, with respect to their assumptions about the nature and purpose of evaluation and the role of the evaluator. It also reflects their perspective on the inter-paradigmatic context of the public debate under study. This group in particular does not seem to engage with the inter-paradigmatic dimension of this context at all. As these NPOs do not question the foundational assumptions of the CMF and engage with it based on the same assumptions, they can be regarded as *monist* – in that they tend to assume a foundational identity between their own perspective on evaluation and the CMF. While this perspective may contribute more to facilitating agreement between NPOs and government, it may also reflect limits in entertaining the possibility of multiple, alternative paradigms of evaluation.

On the other hand, G2 NPOs tend to display features of the interpretivist paradigm, in that they do regard the evaluation process as value laden, therefore accept the plurality of truths and loci of expertise, and the differences in purpose of evaluation between the macro and micro levels. This is reflected in the way they identify foundational differences between their perspective and the CMF, and place the possibility of measurement and centralization at macro level while denying the possibility of planning and design of social interventions due to the unpredictability of outcomes at micro level. Typically interpretivist is also their view that both types of evaluation are justified at their respective levels but not otherwise. Therefore, the solutions proposed by G2 NPOs rely on the idea of a separation between evaluation by government and evaluation by

NPOs into two non-overlapping areas. We label this group as *impartial pluralists*. While this approach may facilitate a form of agreement between NPOs and government, it may also discourage the identification of common goals or mutual engagement in foundational issues about evaluation.

Finally, G3 NPOs share the view that aggregation, centralization and design or planning (understood as fixation on pre-set objectives) are not desirable at any level, as they do not acknowledge the diversity of values involved, the fluid and dynamic nature of local contexts, the unpredictability of complex social outcomes. Most importantly, they are not conducive to meaningful social change that promotes justice and inclusion through the protection and empowerment of the vulnerable. Consequently, G3 NPOs not only critique the foundational assumptions of the CMF by displaying their different paradigmatic orientation, they also insist that their paradigm is more legitimate than the one proposed through the CMF. They are therefore *radical pluralists*, in that they engage with the alternative paradigm in order to gain ground through substantive change – whether by profoundly reforming the proposed evaluation instrument to reflect the more legitimate paradigm (G3a = reformers), or by replacing it with a more legitimate instrument (G3b = displacers/revolutionaries), or by advocating that the measurement and centralization project should be abandoned altogether (G3c = nihilists). While potentially representing the most progressive approach to social change, the position of G3 NPOs may also be the most difficult to deal with in brokering agreement with other social actors such as the government.

## IMPLICATIONS

Based on analysis of empirical data, we have identified that foundational assumptions inform not only normative positions, as aptly illustrated by meta-perspectives on evaluation available in the theoretical literature – but also different perspectives on the inter-paradigmatic contexts organizations like NPOs may find themselves in, and on how tensions between alternative paradigms might be resolved. These findings have a number of implications, for both theory and practice, which we outline in turn below.

### *Implications for Theory*

Building on the meta-perspectives on evaluation offered by Guba and Lincoln (1989), Mertens (2009) and Hall (2014), we extend our understanding of normative positions on evaluation espoused by organizations like NPOs beyond verifying explicit or implicit connections made by NPOs between their normative positions and their foundational assumptions. We further illustrate how different foundational-normative connections inform NPOs' different understandings of inter-paradigmatic contexts and their different approaches to dealing with alternative paradigms. The parallels identified in Table 4 among specific foundational-normative connections, theoretical paradigms of evaluation and perspectives on the inter-paradigmatic context can form the basis for a broader conceptual framework of normative positions on evaluation.

### *Implications for Practice*

By analyzing the public submissions of NPOs as outlined in this study, we seek to improve our understanding of normative positions emerging in inter-paradigmatic contexts in practice by identifying foundational assumptions with the support of insights from evaluation theory.

Furthermore, we have learnt that those NPOs that express clearer and louder disagreement with a

sector-wide evaluation proposal are likely to also be more explicit in defining their paradigmatic differences from the proposal and more aware of the complexities of the inter-paradigmatic context they experience.

While the existence of continuing public disagreements on the issue of evaluation should not be readily assumed to be an undesirable state of affairs, it is essential for the advancement of inter-sector relations, especially in the context of government-NPO collaboration, to expand our understanding of the sources of such disagreements (Hall, 2014, pp. 331-332). Both government and the nonprofit sector may find it useful to understand the role of NPOs' foundational-normative reasoning on evaluation in reacting to social actors with a different paradigm. For both sides, this may lead to increased tolerance for paradigmatic differences in NPO assessment and increased chances for constructive inter-sector communication.

### **SUGGESTIONS FOR FUTURE RESEARCH**

We recommend that current meta-perspectives on evaluation be refined to reflect more comprehensively a variety of logical links that can be established between particular paradigms and particular perspectives on inter-paradigmatic contexts. There is more to be explored, for example, about how value neutral assumptions of evaluation processes lead to insensitivity to inter-paradigmatic situations and alternative paradigms, while value sensitive assumptions lead to awareness of such situations and paradigms. Similarly, there should be more inquiry into why and how the value relativism characteristic of interpretivist positions may lead to less engagement with alternative paradigms, while the transformative focus of constructivist positions may lead to more engagement alternative paradigms but also to less social agreement.

We are also aware that our approach has left aside the political dimension of inter-sector communication. As power relations constitute an important aspect of such communication, we suggest that a comprehensive political theory of paradigm interactions around evaluation should be developed. Insights from Laclau and Mouffe (2001), for example, into the complex relationships between micro-power and hegemony may assist further exploration of the interplay between paradigmatic differences and political interactions in evaluation approaches.

## **CONCLUSION**

Using the 2009-2010 public debate on evaluation as a case study for an examination of paradigm interactions, we identified and analyzed three foundational arguments which then helped us establish a range of connections between different foundational premises and normative positions expressed by NPOs on issues of evaluation. Our analysis highlights the ways in which particular foundational-normative connections reflect particular evaluation paradigms as well as particular perspectives on the inter-paradigmatic context of the NPO sector. We have found that the inter-paradigmatic context illustrated by this debate contributes to the emergence of a variety of normative positions on both how evaluation should be conducted and how alternative paradigms of evaluation should be responded to. While our study constitutes a starting point for understanding the impacts of paradigmatic differences in the theory and practice of NPO evaluation, we call for further study of meta-perspectives on evaluation to better explain such differences.

## **REFERENCES**



Alkin, M. C. (2004), *Evaluation Roots: Tracing Theorists' Views and Influences* (Sage Publications, Thousand Oaks CA).

Alkin, M. C. (2012), *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences* (2<sup>nd</sup> ed., Sage Publications, Thousand Oaks CA).

Brown, E. and K. Gaughlin (2009), 'Donors, Ideologues, and Bureaucrats: Government Objectives and the Performance of the Nonprofit Sector', *Financial Accountability & Management*, Vol. 25 No 1, pp. 99-114.

Carman, J. G. (2007), 'Evaluation Practice among Community-based Organizations: Research into the Reality', *American Journal of Evaluation*, Vol. 28, pp. 60-75.

Carman, J. G. and K. A. Fredericks (2008), 'Nonprofits and evaluation: Empirical evidence from the field', *New Directions for Evaluation*, Vol. 119, pp. 51-71.

Cheverton, J. (2007), 'Holding Our Own: Value and Performance in NPOs', *Australian Journal of Social Issues*, Vol. 2, No 3, pp. 427-436.

Commonwealth of Australia (2001), *Report of the Inquiry into the Definition of Charities and Related Organisations* (Commonwealth of Australia, Canberra).

Creswell J. W. (2013), *Qualitative inquiry and research design: Choosing among five approaches* (Sage, Thousand Oaks CA).

Dart, R. (2004), 'Being "business-like" in a Non-profit Organization: A Grounded Inductive Typology', *Nonprofit and Voluntary Sector Quarterly*, Vol. 33, No 2, pp.290-310.

Ebrahim, A. and E. Weisband (eds.) (2007), *Global Accountabilities: Participation, Pluralism, and Public Ethics* (Cambridge University Press, Cambridge).

Eckerd, A. and S. Moulton (2011), 'Heterogeneous Roles and Heterogeneous Practices: Understanding the Adoption and Uses of Nonprofit Performance Evaluations', *American Journal of Evaluation*, Vol. 32, pp. 98-117.

Frumkin, P. (2002), *On Being Nonprofit: A Conceptual Policy Primer* (Harvard University Press, Cambridge MA).

Guba, E. G. and Y. S. Lincoln (1989), *Fourth Generation Evaluation* (Sage Publications, Newbury Park CA).

Guba, E. G. and Y. S. Lincoln (2005), 'Paradigmatic Controversies, Contradictions, and Emerging Confluence', in N. K. Denzin and Y. S. Lincoln, *The SAGE Handbook of Qualitative Research* (3rd ed., Sage, Thousand Oaks CA).

Gugerty, M. K. (2009), 'Signaling Virtue: Voluntary Accountability Programs among NPOs', *Policy Sciences*, Vol. 42, pp. 243-273.

Hall, M. (2014), 'Evaluation Logics in the Third Sector', *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, Vol. 25, No 2, pp. 307-336.

Herman, R. D. and D. O. Renz (1999), 'Theses in NPO Organizational Effectiveness', *Nonprofit and Voluntary Sector Quarterly*, Vol. 28, No 2, pp. 107-126.

Herman, R. D. and D. O. Renz (2008), 'Advancing Non-profit Organizational Effectiveness Research and Theory: Nine Theses', *Nonprofit Management and Leadership*, Vol. 18, pp. 399-415.

- Housego, A. and T. O'Brien (2012), 'Delivery of Public Services by Non-Government Organizations', *Australian Journal of Public Administration*, Vol. 71, pp. 211-220.
- Hwang, H. and W. W. Powell (2009), 'The Rationalization of Charity: The Influences of Professionalism in the Nonprofit Sector', *Administrative Science Quarterly*, Vol. 54, pp. 268-298.
- Industry Commission (1995), *Charitable Organizations in Australia. Report No 45, 16 June 1995* (Commonwealth of Australia, Canberra).
- Laclau, E. and C. Mouffe (2001), *Hegemony and Socialist Strategy: Towards a Radical Democratic Politics* (2nd ed., Verso, London).
- Lewis, L. (2005), 'The Civil Society Sector: A Review of Critical Issues and Research Agenda for Organizational Communication Scholars', *Management Communication Quarterly*, Vol. 19, No 2, pp. 238-267.
- Lincoln S. Y. and E. G. Guba (1985), *Naturalistic Inquiry* (Sage, Thousand Oaks CA).
- Mertens, D. M. (1999), 'Inclusive Evaluation: Implications of Transformative Theory of Evaluation', *American Journal of Evaluation*, Vol. 20, No 1, pp. 1-14.
- Mertens, D. M. (2009). *Transformative Research and Evaluation* (Guilford Press, New York).
- Moulton, S. and A. Eckerd (2012), 'Preserving the Publicness of the Nonprofit Sector', *Nonprofit and Voluntary Sector Quarterly*, Vol. 14, No 4, pp. 656-685.
- Murray, V. (2005), 'Evaluating the effectiveness of nonprofit organizations', in R. Herman and Associates, *The Jossey Bass Handbook of Nonprofit Leadership and Management* (2nd ed., Jossey-Bass, San Francisco CA).

Owen, J. M. (2006). *Program Evaluation: Forms and Approaches* (Allen & Unwin, St Leonards NSW).

Polonsky, M. and S. Landreth Grau (2008), 'Evaluating the Social Value of Charitable Organizations: A Conceptual Foundation', *Journal of Macromarketing*, Vol. 28, No 2, pp. 30-140.

Productivity Commission (2010), *Contribution of the Not-for-Profit Sector: Productivity Commission Research Report, January 2010* (Commonwealth of Australia, Canberra).

Quarter, J. and B. J. Richmond (2001), 'Accounting for Social Value in NPOs and For- Profits', *Nonprofit Management & Leadership*, Vol. 12, No 1, pp. 75-85.

Scriven, M. (1980), *The Logic of Evaluation* (Edge Press, Port Reyes CA).

Sen, A. (1999), *Development as Freedom* (Oxford University Press, Oxford).

Sowa, J., Selden, S., & Sandfort, J. (2004). 'No longer unmeasurable? A multidimensional integrated model of nonprofit organizational effectiveness', *Nonprofit and Voluntary Sector Quarterly*, vol. 33, No 4, pp. 711-728.

Stufflebeam, D. L. (2001), 'Evaluation Models', *New Directions for Evaluation*, Vol. 85, pp. 8-98.

Thomson, D. E. (2011), 'The Role of Funders in Driving Nonprofit Performance Measurement and Use of Strategic Management', *Public Performance and Management Review*, Vol. 35, No 1, pp. 54-78.

Table 1  
Evaluation paradigms, their foundational assumptions and normative positions

Paradigms	Meta-perspectives on evaluation			Foundational assumptions		Normative position
	Guba and Lincoln's (1989)	Mertens (1999, 2009)	Hall (2014)	Nature of evaluation	Role of evaluator	Purpose of evaluation processes
Positivist	<ul style="list-style-type: none"> <li>▪ Measurement: gauge (technical) phenomena</li> <li>▪ Description: describe strengths and weaknesses</li> <li>▪ Judgment: evaluate between options</li> </ul>	<ul style="list-style-type: none"> <li>▪ Post-positive: discover single objective knowable truth</li> </ul>	<ul style="list-style-type: none"> <li>▪ Scientific: collect and analyze data</li> <li>▪ Bureaucratic: categorize and assess effects</li> </ul>	Evaluation = technical measurement process; stable entity; reflects essential/evident truth; value neutral	Evaluator = expert/authoritative actor	Measurement as a tool for discovering a single objective truth
Interpretivist		<ul style="list-style-type: none"> <li>▪ Interpretive: create plural, socially constructed realities</li> </ul>	<ul style="list-style-type: none"> <li>▪ Learning: assess, discuss, revise knowledge</li> </ul>	Evaluation = social process of continuous negotiation; discovering “truth”; value sensitive but impartial	Evaluator(s) = multiple actors; plural and subjective experiences	Evaluation processes as a means of reaching consensus about existing phenomena (mostly with the intention to “improve” them)
Constructivist	<ul style="list-style-type: none"> <li>▪ Responsive constructivism: create realities through social and political negotiations</li> </ul>	<ul style="list-style-type: none"> <li>▪ Transformative: challenge and change social conditions</li> </ul>		Evaluation = social and political process; continuous negotiation; plural and subjective realities; value laden	Evaluator(s) = multiple actors, creating “truth”	Evaluation as an arena in which to challenge and transform existing phenomena (e.g. confront social inequality)

Table 2  
Foundational arguments and normative positions in NPO submissions: examples

	Foundational Arguments			Normative Position (relative to CMF)
Sub No	Measurability/ Non-measurability	Centralization/ Contextualization	Design/ Emergence	
#135 (G1)	Can measure: 'Government and the sector need to refocus and develop measurement tools that measure against indicators of scientific, social and cultural capital not just in economic terms.' (p. 8)	Can centralize: 'Government needs to... view effective evaluation as a way to... determine... that services and programs are aligned with strategic directives...' (emphasizes aggregating results as a function of government) (p.9)	Can plan: 'Evaluation should be utilised to inform and provide the evidentiary basis necessary for appropriate strategic planning...' (p. 9)	CMF should be adopted: 'measurements need to be developed that actually measure the impacts of the service rather than simply measuring economics and counting the numbers of service users.' (p. 9) – identical with Report recommendation (CMF proposal) at p. xxxv.
#165 (G2)	Can measure (macro level): Example of useful measure at macro level: 'Not for profit contributions could be assessed against key social sustainability indicator that could be included in a national charter' (p. 3)	Can centralize (macro level): [Uniform measurement standards] 'could... enable the production of consistent and comparable sector wide data'. (p. 5)	Cannot plan: [Pre-setting macro objectives does not work because] '... the diversity of providers ensures a tailored response that addresses the diversity of client needs' (p. 6)	CMF should be adopted <u>only</u> at macro level: 'Red Cross commends... a robust framework...' (p. 1) 'To support... meaningful indicators... and enable comparisons and... [government] use of the data to inform [national] policy... specific and additional funds need to be provided' (p. 3) (evaluation at macro and micro level are two separate exercises, which require separate resources)
#140 (G3a)	Cannot measure: '...numbers alone will tell us nothing about the quality of ...participation...' (p. 7)	Cannot centralize: 'The experience of the Anglicare network indicates the need for considerable vigilance in making generalizations.' (p. 8)	Cannot plan: [Effective service delivery means] 'being open [flexible] about the... underlying objectives.' (p. 11)	CMF should be significantly modified: 'the conceptual framework suggested by the Commission should be integrated with an appropriate variant of the capabilities approach developed by Sen' (p. 13)
#138 (G3b)	Cannot measure: [A quantitative] 'approach has limited success in addressing a person's holistic needs' (p. 15)	Cannot centralize: 'Government funding models are function/silo based rather than person/community focussed.' (p. 13)	Cannot plan: 'Many organisations in the NFP sector ... develop organically to meet changing needs ... of... service users.' (p. 8)	CMF should be replaced: '[Government should follow NPO practices and] develop people and community focussed... frameworks... to implement an integrated approach to poverty alleviation and social inclusion.' (p. 14)
#117 (G3c)	Cannot measure: 'Much of the sector's contribution ...[is] incommensurable with ...social efficiency'(p.4)	Cannot centralize: 'Not all relationships can be valued according to the same metric.' (p. 12)	Cannot plan: 'It is a mistake to think that policy makers [should] always begin with a fully developed set of objectives.' (p. 8)	CMF should not be used at all: 'What if government had to choose between measuring the contribution of the not-for-profit sector and facilitating that contribution?' (p. 4)

Table 3

Foundational arguments and normative positions in NPO submissions: NPO group composition

	Foundational Arguments			Normative Position (relative to CMF)	NPOs (n = 54)
	Measurability/ Non-measurability	Centralization/ Contextualization	Design/ Emergence		
G1	Can measure	Can centralize	Can plan	CMF should be adopted	Ability Options (119), Alcohol and Other Drugs Council of Australia (123, 149, DR267 <sup>8</sup> ), Centre for Excellence in Child and Family Welfare (DR286), Mission Australia (56, DR220), Network of Alcohol and Drug Agencies (66, DR196), PilchConnect (131, DR277), Public Interest Advocacy Centre (174), South Australian Council of Social Service (135), The Australian Lung Foundation (49), Women's Health Victoria (DR294) (n = 10)
G2	Can measure (macro level)	Can centralize (macro level)	Cannot plan	CMF should be adopted as a separate evaluation exercise at macro level, without interference with evaluation approaches at micro level	Aged and Community Services Australia (86), Australian Council of Social Service (118, DR256), Australian Institute of Company Directors (DR239), Australian Red Cross (165, DR296), Communities@Work (150), Fundraising Institute Australia (76, DR222), National Disability Services (85, DR263), Philanthropy Australia (62, DR253), Queensland Aged and Disability Advocacy Inc. (103), Social Traders (102, DR189), The Benevolent Society (100, DR225), The Myer Foundation (128), The Wilderness Society (DR282), Victorian Council of Social Service (164, DR276), Vision Australia (DR227), Yooralla (92) (n = 16)

<sup>8</sup> DR stands for 'Draft Report' and indicates a second submission by the NPO in response to the Commission's Draft Report.

G3	Cannot measure	Cannot centralize	Cannot plan	CMF should not be adopted in its current form:	
G3a				CMF should be significantly modified to align with micro level evaluation priorities	Anglicare Australia (140), Association of Neighbourhood Houses and Learning Centres (122, DR303), Australian Research Alliance for Children and Youth (DR199), Australian Library and Information Association (124), Ethnic Communities Council of Victoria (60), Illawarra Forum (52, DR232), Independent Schools Council of Australia (DR278), National Association of People Living with HIV (DR300), National Breast Cancer Foundation (98), The Federation of Ethnic Communities Councils of Australia (61), The Smith Family (59, DR204) (n = 12)
G3b				CMF should be replaced by the more widespread approaches already practised by the NPOs at micro level	Australian Council for International Development (136, DR299), Cancer Council Australia (DR318), Community and Neighbourhood Houses Association (25), Friends of Libraries Australia (14), Kids Under Cover (DR250), New South Wales Family Services (FamS) (158), UnitingCare Australia (138, DR291) (n = 7)
G3c				G3c: CMF should not be used at all, as macro level evaluation is not meaningful	Australian General Practice Network (151), Berry Street (51, DR283), Catholic Social Services Australia (117), Family Relationship Services Australia (132), Health and Community Services Workforce Council (95), Local Community Services Association (144), Peak Care Queensland (81), Room to Read (DR228), Social Innovation, Western Australian Network for Alcohol and Other Drug Agencies (137) (n = 9)



Table 4  
Paradigmatic orientations and perspectives on the inter-paradigmatic context in NPOs

NPO Group	Paradigmatic Orientation	Perspective on Inter-paradigmatic Context
G1	<p><i>Positivist</i></p> <ul style="list-style-type: none"> <li>- Confidence in value neutral processes: <u>confidence in measurement</u><sup>9</sup></li> <li>- Single locus of expertise: <u>centralization is justified</u></li> <li>- Single objective truth accessible to expert: <u>confidence in planning</u></li> </ul>	<p><i>Monist</i></p> <ul style="list-style-type: none"> <li>- NPO and CMF share the same paradigm / there are no paradigmatic differences between the NPO perspective on evaluation and the CMF</li> </ul> <p>Outcome = unproblematic agreement</p>
G2	<p><i>Interpretivist</i></p> <ul style="list-style-type: none"> <li>- More than one truth: <u>measurement and centralization (macro level) vs non-measurement and contextualization (micro level)</u></li> <li>- More than one locus of expertise: <u>cannot plan both levels together, in the same way</u></li> <li>- Value sensitive process – but impartial: <u>both evaluation types are justified, at different levels</u></li> </ul>	<p><i>Impartial Pluralist</i></p> <ul style="list-style-type: none"> <li>- NPO perspective and CMF rely on different paradigms, <u>and</u></li> <li>- these paradigms operate at different levels of evaluation (micro level vs macro level), <u>and</u></li> <li>- both paradigms are equally legitimate</li> </ul> <p>Outcome = separation into two non-overlapping (mutually exclusive) areas of evaluation / seeking agreement on mutual acknowledgment and non-interference</p>
G3	<p><i>Constructivist</i></p> <ul style="list-style-type: none"> <li>- Value sensitive process: <u>cannot measure different subjective values</u></li> <li>- Mistrust of ‘expertise’: <u>cannot centralize local knowledge (dynamic); cannot plan complex social outcomes (unpredictable)</u></li> <li>- ‘Truth’ legitimacy related to protection and empowerment of the vulnerable: <u>focus on social change, justice and inclusion</u></li> </ul>	<p><i>Radical Pluralist</i></p> <ul style="list-style-type: none"> <li>- NPO perspective and CMF rely on different paradigms, <u>and</u></li> <li>- these paradigms operate at different levels of evaluation (micro level vs macro level), <u>and</u></li> <li>- the micro level is more legitimate than the macro level</li> </ul> <p>Outcome = seeking assertion of superior legitimacy of own paradigm:</p> <ul style="list-style-type: none"> <li>- profound modification (G3a = reformers)</li> <li>- replacement (G3b = revolutionaries)</li> <li>- abandonment (G3c = nihilists)</li> </ul>

<sup>9</sup> The underlined text in the second column represents the specific ideas espoused by the NPOs in their submissions.